

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-306414

(43)Date of publication of application : 02.11.2001

(51)Int.Cl.

G06F 13/00

G06F 3/06

G06F 13/12

H04L 13/08

(21)Application number : 2000-128970

(71)Applicant : HITACHI LTD

(22)Date of filing : 25.04.2000

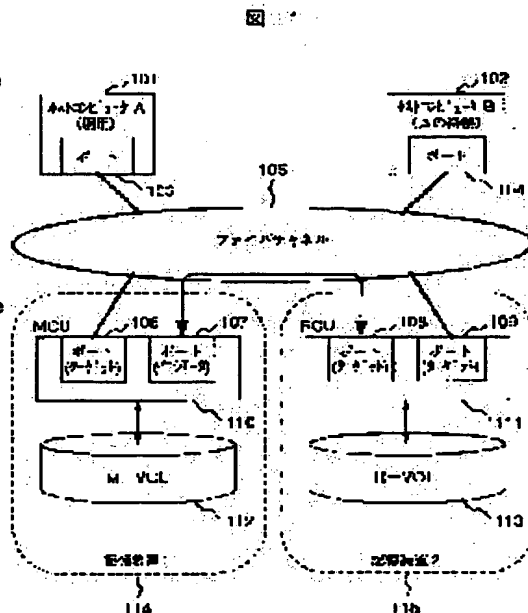
(72)Inventor : KOBAYASHI NAOTAKA
ABEI MASARU

(54) REMOTE COPYING SYSTEM FOR STORAGE DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To solve such problems that the remote copy transfer speed of a storage device is slow in a compute system for realizing remote copying between plural storage devices and thereby the throughput of a host I/O has been interrupted.

SOLUTION: A host computer is connected to plural storage devices through a fiber channel and remote copying is performed through the fiber channel. Thereby a storage device to be a copying source performs log-in by applying information capable of identifying log-in from the storage device itself and a storage device for receiving the log-in determines a remote side storage device and a port candidate by returning information for specifying a port capable of performing remote copying only at the time of receiving the log-in from the storage device concerned.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim (a)]

[Claim 1] Storage with the port used as the initiator of a host computer and a remote copy and storage with a target port are connected through a fiber channel. As for the storage with the port used as the aforementioned initiator, self has the control unit which performs login which added the 1st information which can discriminate that it is storage in the target port of storage with the aforementioned target port. When, as for the storage with the aforementioned target port, a logging in agency discriminates that it is storage using the 1st information of the above on the aforementioned login. The remote copy system of the storage characterized by having the control unit which returns the response which added the 2nd external information peculiar to the hardware which pinpoints the aforementioned target port and.

[Claim 2] The 2nd information of the above is the remote copy system of the storage according to claim 1 characterized by being the identifier peculiar to hardware and port watch of storage with a target port.

[Claim 3] The 2nd information of the above is the remote copy system of the storage according to claim 1 characterized by being **** and port watch with a target port of storage.

[Claim 4] The 1st information of the above is the remote copy system of the storage according to claim 1 characterized by being the identifier peculiar to hardware and port watch of storage with the port used as an initiator.

[Claim 5] The 1st information of the above is the remote copy system of the storage according to claim 1 characterized by being **** and port watch with the port used as an initiator of storage.

[Claim 6] The aforementioned port watch is the remote copy system of any of the claims 3, 4, and 5 characterized by being the identification number of the location which shows the wearing position of the port held with the storage used as an initiator, or the storage of a publication.

[Claim 7] The storage which is the copy origin of a host computer and a remote copy, and has a master control unit, and storage with a target port are connected through a fiber channel. From an initiator port, the aforementioned master control unit adds the information which shows that it is login from a master control unit to a target port, and logs in. As storage with which self will be set as the object of a remote copy if it recognizes that the storage with the aforementioned target port is login from a master control unit. The remote copy system of the storage characterized by returning the response which added external information peculiar to the hardware which pinpoints the aforementioned target port and to the aforementioned initiator.

[Claim 8] The information which shows that it is login from the aforementioned master control unit is the remote copy system of the storage according to claim 7 characterized by being an identifier peculiar to the aforementioned master control unit, and port watch.

[Claim 9] The aforementioned port watch is the remote copy system of the storage according to claim 8 characterized by being the identification information of the location which shows the wearing position of the port held by the aforementioned master control unit.

[Claim 10] The information which shows that it is login from the aforementioned master control unit is the remote copy system of the storage according to claim 7 characterized by being identifiers other than the identifier used by the fiber channel communications protocol.

[Claim 11] Two or more storage which has a control unit, respectively is connected with a host computer through a fiber channel. The control unit of the storage which becomes the copy origin of a remote copy attaches an identifier peculiar to the control unit concerned, and port watch, and logs in to the port of other aforementioned storage. It answers that the control unit of storage besides the above has the identification number peculiar to a control unit and port watch of the storage which becomes data of the aforementioned login the copy origin of a remote copy. The remote copy system of the storage characterized by returning the response which added an identifier peculiar to the control unit concerned, and port watch.

[Claim 12] The remote copy system of storage characterized by providing the following. The control unit of the storage which becomes the copy origin of the remote copy which two or more storage which has a control unit, respectively is connected with a host computer through a fiber channel, attaches the information which can distinguish that it is login from storage, and logs in to the port of other aforementioned storage. The control unit of storage besides the above which returns the response which answered that the information which can distinguish that it is login from the aforementioned storage was in the data of the aforementioned login, and added an identifier peculiar to the control unit concerned, and port watch. The target managed table which stores an identifier peculiar to the aforementioned control unit contained in a response, port watch, and the destination address of the port. The logical-path-target translation table which stores the target port which has an identifier peculiar to the aforementioned control unit in common, and port watch in an identifier peculiar to the control unit of the copy place specified by the application layer, the logical-path managed table which stores port watch, and the aforementioned target managed table and the aforementioned logical-path managed table as a port of a copy place.

[Claim 13] Port watch peculiar to the control unit of the storage which becomes the aforementioned copy origin is the remote copy system of any of the claims 11 and 12 characterized by being the identification information of the location which shows the wearing position of the port held by the control unit of the storage concerned, or the storage of a publication.

[Claim 14] The remote copy system of any of the claims 11 and 12 characterized by a port number peculiar to the control unit of other storage different from the aforementioned copy origin being the identification information of the location which shows the wearing position of the port held by the control unit of the storage concerned, or the storage of a publication.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. *** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] this invention relates to the remote copy between the storage of a computer system.

[0002]

[Description of the Prior Art] Creating the copy of the volume (stored data) of storage, such as disk storage of a computer system, to another storage in a remote place is performed. This copies with it, when it becomes the situation which cannot read the data of the storage of a copied material by generating of calamities, such as an earthquake, or a certain serious obstacle, and it is used instead of the data of copy place copy origin at this time. Creating such a copy or the created copy is called remote copy. The remote copy is adopted as the system by which high reliability is demanded especially, for example, a banking system etc.

[0003] What made full use of ESCON technology exists in the remote copy between the conventional storage. JP.6-236340A is mentioned as the well-known example.

[0004] By the way, the fiber channel is used for the transfer of the information between a host computer and storage as a channel interface for fast transmission in recent years. In the protocol of such a fiber channel, in order to discriminate each equipment (node) with which it is dotted through a fiber channel, or its port from others, the unique information based on the protocol specification is used. WWN (World Wide Name), AL_PA (Arbitrated Loop Physical Address), etc. are so.

[0005] When a fiber channel is used for a data communication means, there are two big merits, the performance with a quick (a maximum of 100 MB/S) transmission speed in which a cable can be extended for a long time (infinite [in a maximum of 10km and fabric connection]) by the simple substance (cable), and distance. Also in the aforementioned storage, this merit is obtained by the system configuration with the host computer using the fiber channel.

[0006]

[Problem(s) to be Solved by the Invention] By the way, in order to require the data transfer from storage to the local storage of RIMOTO in the system which carries out a remote copy, as compared with the case where a remote copy is not carried out, the throughput of host I/O (I/O of the data between a host and storage) falls certainly. It is because the light data which were held on the data cache which storage has in the interior of equipment in the remote copy are left behind on a data cache for a long time than usual, and is because the use efficiency of a data cache is bad. The cause that light data are left behind to a data cache for a long time originates in the transfer performance of a remote copy. ESCON is used for the remote connection of the remote copy to the former as above-mentioned, and although a remote copy is performed even if the transfer rate of a remote copy is very slow and it uses a fiber channel for the transfer means between a host computer and storage as a result compared with the transfer rate of host I/O, the transfer performance of host I/O will be decreased to a sake.

[0007] Although it is possible in one more point and ESCON to extend a cable further by going via an extender in long distance connection (10km **), the device of an ESCON system is very expensive. Even when it goes via a fabric switch or a hub by the point fiber channel, if compared

with the device of an ESCON system, since it is very cheap, also in the cost side on a system construction, it is large to the expensive *****.

[0008]

[Means for Solving the Problem] From the above thing, this invention is using a fiber channel for the data transfer of a remote copy, and hindrance of the throughput of host I/O will be made into the minimum. In order to realize the above, in this invention, it had the initiator port which can communicate through the fiber channel on the storage used as the source, the storage of a remote site is equipped with the target port connected to a fiber channel, and data transfer is performed through the fiber channel protocol for the linkup in a fiber channel. Furthermore, from an initiator port, the information which can recognize that self-equipment is storage is attached, and it logs in to a target port, and when it has recognized that login issue origin is storage, a target port attaches the eternal hardware information which pinpoints a self-port, and returns it to an initiator port.

[0009]

[Embodiments of the Invention] Hereafter, the example of this invention is explained using a drawing. Drawing 1 shows the remote copy structure of a system used as the example of this invention. Host computers 101-102 all have a port for fiber channel interfaces (respectively 103-104), and are physically connected with storage 114-115 through the fiber channel interface. The storage 114-115 as well as host computers 101-102 has the port for fiber channel interfaces (respectively 106-107, 108-109), and communication by the fiber channel protocol is possible for it. Although some kinds, such as Point-to-Point, and Arbitration loop connection, fabric connection, exist in the topology (topology) of the fiber channel interface between host computers 101-102 and storage 114-115, this invention is only described to be the fiber channel 105, in order not to be dependent on the topology.

[0010] It aims at copying to the data volume (R-VOL113) in which the storage 115 of RIMOTO has the data volume (M-VOL112) which storage 114 has through the fiber channel 105 in this remote copy system. The host interface control unit of the storage 114 used as a master calls the control unit which has the port 107 used as an initiator and has the port 108 which calls this control unit master control unit (MCU) 110, and serves as a target of the receiving remote copy place remote control unit (RCU) 111 in that case. In this invention, it is lifting with the first feature using a fiber channel for the interface protocol used between MCU-RCU. Then, the outline of the protocol will be explained.

[0011] A fiber channel is a protocol with the serial transmittal mode without an original commands set, and in order to send information asynchronously, it has the special feature which can use the bandwidth of a transmission medium effectively. Reliable colorful data transfer is more possible at high speed, inheriting the conventional protocol property by using a fiber channel for instead of [without an original commands set] as an infrastructure of commands sets, such as the conventional SCSI, ESCON, HIPPI, IP-3, and IP.

[0012] A fiber channel is an interface which has a channel and the feature of a network. Although delay can realize a high-speed few transfer by the fiber channel once the source and the destination are decided, this is one of the greatest features of a channel. Moreover, this is the feature of a network, although the device which wishes to communicate can participate in the communication system of a fiber channel by arbitrary opportunities, the agreement information about communication can be exchanged for the device of the partner who becomes the communicative purpose, and mutual and communication can be started. The procedure of the agreement information interchange about communication with the partner device described here is called login.

[0013] A device (for example, a host computer and storage) with the interface of a fiber channel is called node, and the physical mouth which hits an actual interface is called port. A node can have one or more ports. The number of nodes which can participate in the whole system of a fiber channel simultaneously is the number of the addresses of 24 bits, i.e., about 16,770,000 pieces, at the maximum. Generally, the fiber channel system consists of a fiber channel of the shape of a lot of loop, connects between loops, and calls fabric the hardware which relays information. In fact, a transmitting agency and the port of a transmission place should just

operate only in consideration of the information about a mutual port, without being conscious of a fabric. Drawing 1 is also simplified and expressed. The identifier unique all over the world assigned by the fixed rule from a standardization organization (IEEE) is memorized in each node and the port. This is equivalent to a familiar Media Access Control Address by TCP/IP etc. from the former, and is the address [**** / in hardware]. There are two kinds of these addresses, N_Port_Name and Node_Name, N_Port_Name is a value (hardware address) peculiar to a port, and Node_Name is a value (hardware address) peculiar to a node. Since each of these is values unique all over the world, they are called WWN (World Wide Name) as the address which can discriminate a node or a port uniquely.

[0014] Communication is performed by the fiber channel for the information on the signal level called Ordered Set, and logical information with the format of the fixation called frame. Drawing 2 shows the structure of a frame. A frame 201 is twisted from 4 bytes of identifier called the round [of 204 or 4 bytes of data field which is a part for data division used as 4 bytes of identifier called SOF (Start of Frame)202 which shows the beginning of a frame, 24 bytes of frame header 203 which performs control of link motion, and characterization of a frame, and the purpose actually transmitted] redundant code (CRC) 205, and EOF (End of Frame)206 which shows the end of a frame. A data field 204 is adjustable among 0~2112 bytes.

[0015] Next, the login procedure based on a fiber channel protocol which exchanges information mutually about the device of a transmitting agency and the device of a transmission place is described. The structure of the PLOGI (port login) frame which becomes indispensable, and the PRLI (process login) frame is explained in the case of login procedure. In the detailed structure 207 of PLOGI of a data field 204, 8 bytes of field from a head to 21st byte - the 29th byte is a field which stores N_Port_Name208, and 8 bytes of field from a head to 30th byte - the 38th byte stores Node_Name209. Moreover, in the detailed structure 210 of PRLI, the 4-byte field from a head to 8th byte - the 11th byte is a field which stores the parameter of cage JINEITA process ASOSHIEITA (Originator process associator) 211. This is parameter area which becomes effective, when using a node as an initiator. Moreover, the 4-byte field from a head to 12th byte - the 15th byte is a field which stores the parameter of responder process ASOSHIEITA (Responder process associator) 212. This is parameter area which becomes effective, when using a node as a target.

[0016] Drawing 3 shows the exchange of the information exchanged between 301 and the transmission place (login reception place) 302 a transmitting agency (login demand origin). Although it exists in the login procedure of a fiber channel in some numbers, here describes login of a class 3. A class 3 does not show one of the types of the communication procedure of a fiber channel, and especially application of this invention is not restricted to this. It illustrates on account of explanation.

[0017] Login demand origin transmits the PLOGI frame 303 to a login reception place. N_Port_Name of login demand origin, Node Name, and the information on other are included in this frame. With the equipment of a reception place, the information included in this frame is taken out, and when recognizing login, the frame called ACC304 is transmitted to login demand origin. On the other hand, when refusing login, the frame called LS_RJT305 is transmitted to login demand origin. It is not decided [which recognizes login / or] in the light of the rule of the fiber channel which a target side has whether to carry out refusal, and it is not directly related to this invention. Login demand origin will know that login was successful, if the response of the ACC frame is detected to the PLOGI frame which oneself transmitted. A success of login between ports performs an exchange on process level continuously. Login demand origin transmits the PRLI frame 306 to a login reception place. The information which exchanges the operating environment of the FCP (fiber channel protocol) level between the process groups related in two ports is included in this frame. With the equipment of a reception place, like the time of PLOGI receipt, when recognizing login, when refusing, the LS_RJT frame 308 is transmitted to login demand origin to ACC307. If the login procedure so far is successful, it will be in the state where I/O processes, such as data transfer, can be started. It is a general procedure between ports when the above uses a fiber channel.

[0018] The remote copy which used the fiber channel for below is explained. In drawing 4, the

exchange through the fiber channel between MCU401-RCU402 is shown. By MCU401-RCU402, login procedure for establishing the path on a fiber channel is performed as above-mentioned. First, MCU401 to PLOGI is transmitted and ACC which notifies recognition of PLOGI is transmitted in RCU402. There is no change in any way with the case where the target device in which PLOGI was sent from the host and this received it carries out an ACC response. Then, although MCU401 transmits PRLI in this invention, as a means for telling to RCU that the storage concerned is MCU401, it puts **** (it can be said also as **** of the storage equipped with MCU) and port watch of MCU on the 2nd (parameter field of Originator Process associator) W of the pay load (live data of a data field) 403 of the PRLI frame, and transmits to it. RCU402 which received the PRLI frame starts the 2ndW of the pay load of a frame, and takes out **** and port watch of MCU. Thereby, RCU can discriminate that the initiator concerned is MCU. **** is an identification number unique (it is peculiar to hardware) to hardware, and it is eternal irrespective of change of a fiber channel system, exchange of a package, etc. If usually soft, the thing which is called **** and which is not recognized is embedded into transmit data, and recognition is made possible. Here, it has the implications that it is the peculiar identification information that it is storage turns out to be. In this example, although **** was used, as long as it is the information which can recognize that it is otherwise storage, you may be identifiers other than the identifier (identifier used by the communications protocol of a fiber channel) on the specification currently used, other information, for example, fiber channel. Thus, RCU distinguishes login from MCU, and login from a host computer by the existence of **** and port watch. Of the one storage, MCU port watch is peculiar and eternal.

[0019] Port watch is identification information which shows the wearing position in the storage of a port provided in storage, and it defines as a peculiar location number within one storage. For example, as long as it is the storage in which a maximum of 32 port possession is possible, you may give the numbers from one to 32. Even if this port watch equips with a port substrate with another N_Port_Name since it is information which shows a location for example, it is eternal. On the other hand, the identification information shown by N_Port_Name is the address given to each port, and is dependent on the hardware (for example, port substrate) which constitutes a port. Therefore, a port substrate breaks down, and if it equips with another port substrate, the address of the port concerned will be changed.

[0020] If it proves that the initiator which has transmitted PRLI is MCU, RCU will put the watch made from RCU, and RCU port watch on the 3rd (parameter field of Responder Process associator)W of the pay load 404 of the ACC frame, and will answer it. MCU which received the ACC frame starts the 3rdW of the pay load of a frame, and takes out **** and port watch of RCU. Thereby, MCU can pinpoint the port of RCU while being able to recognize that it is one of the storage with which the partner who logged in can do a remote copy. Thus, the flow chart which divided into the receipt place the flow of the processing which incorporated the method which discriminates a target port in the usual login procedure the requiring agency by drawing 5 and drawing 6 has explained in detail.

[0021] Since it aims at extracting only RCU with which it is finally dotted on a fiber channel in the login procedure which MCU performs as shown in the processing flow of the demand origin of drawing 5, By the ACC response to PRLI after PLOGI being successful, the watch made from an ACC frame shell. Only the target port which was able to take out port watch has a meaning as an object of a remote copy. LOGO (logout) is published to the other target port, and logout processing is performed. This login procedure is performed to all the target ports on a fiber channel, and the target port which was able to take out **** and port watch serves as a candidate of RCU which forms the pair which actually performs a remote copy (it depends on specification from which storage a remote copy is actually performed to which storage of the application program of a high order).

[0022] Order is explained for drawing 5 later on. As shown in drawing 3, PLOGI is published to a certain target port (501). Issue of PLOGI judges whether it is a success (502). Namely, LS if RJT is received, it will move to login failure processing. If ACC is received, AL_PA of the login place contained in it and WWN are acquired, and the partner point is specified (503). Next, PRLI is published (504). And it judges whether it is a PRLI issue success (505). LS if RJT is received, it

will move to login failure processing. Receipt of ACC performs processing which takes out the watch made from a pay-load shell included in it, and port watch (506). Next, ejection judges whether it is a success (507). Since it is that a partner says that it is the storage which can carry out a remote copy, saying that ejection was successful mentions as a candidate of a remote copy place. Since it is shown that a partner is except storage, saying that ejection is impossible carries out logout processing (508). This is continuously repeated to all target ports as mentioned above.

[0023] On the other hand, drawing 6 is the processing flow seen from the receipt place of login. At a receipt place, if the content of a frame is checked when the login frame which a requiring agency transmits is received, it judges [which recognizes login / or or] whether refusal is carried out. With Book RCU, only when the initiator of a requiring agency is MCU, it notifies to MCU that self is RCU, and when that is not right (usually login demand from a host), it answers as a mere target port. The processing for it is the processing which takes out the watch made from MCU, and port watch from the PRLI frame, and only when it is able to take out, own (RCU) **** and port watch are put on the ACC frame, and it answers. It is set to MCU which can form the pair in which this login response is performed to all the initiator ports on a fiber channel, and the initiator which returned **** and port watch actually performs a remote copy.

[0024] Order is explained for drawing 6 later on. PLOGI is received at a receipt place (601). And it judges whether the frame is checked (602) and port login is recognized (603). It is LS when not recognizing. RJT is transmitted and it moves to login (608) failure processing. When recognizing, ACC is created and it transmits (604). If PRLI is received (605), a frame check will be carried out (606) and it will judge whether process login is recognized (607). When not recognizing, it is LS similarly. RJT is transmitted (608) and it moves to login failure processing. When recognizing, extraction processing of **** of a transmitting agency and port watch is performed from a pay load (609). If it succeeds in ejection (610), since this shows that an initiator is MCU, it will set own-making watch and port watch to the ACC frame (611), and will transmit this (612). ACC is transmitted, without setting **** and port watch, since this is login from nodes other than MCU, if it does not succeed in ejection (612).

[0025] At this time, even if it can perform the usual communication, it is not set as the object of a remote copy.

[0026] It is reasonable in MCU having used **** and port watch by the side of RCU for the means for discriminating RCU as mentioned above. By the fiber channel, there is WWN previously described as the address for discriminating the port of a meaning, and AL_PA which appears in a frame as discernment ID of all frames exists in ANSI. Usually, it is common to port discernment on a fiber channel to use these addresses. WWN is information available at the time of the usual login (PLOGI) processing, and AL_PA is information available at the time of the usual port initialization. Two kinds of this identification information is unique on a fiber channel, and the target concerned cannot become the information which shows RCU and there is also no guarantee with the eternal value only at it. Since it depends hard, WWN has a possibility of changing, at the time of hard exchange, and one AL_PA may change it at the time of loop initialization, when a port with the value same on a loop exists. The point, the watch made from equipment, and port watch are the external information independent of a fiber channel, and the logic which carries and returns this to the PRLI frame can be called method effective in RCU discernment, and optimal. Although it specifies and this example explains that it was called the 2nd word of the PRLI frame pay load 403, and the 3rd word of the ACC frame pay load 404 to the storing position of identification information, there is no reason which limits a pay-load position, and it only specifies that it is the method which can pinpoint a target port to the last except information peculiar to fiber channels, such as WWN and AL_PA.

[0027] Thus, MCU can decide the port which serves as RCU out of the target port with which it is dotted on a fiber channel. In MCU, in order to manage RCU detected on the fiber channel, the table shown in drawing 7 is used. The Target managed table 701 registers into PLOGI in order AL_PA taken out from the ACC frame of the target which has carried out the ACC response, and WWN. Then, the watch made from a target chisel with effective Responder Process associator and port watch of the ACC frame are taken out from the target which has carried out

the ACC response also to PRLI, and the term of **** of the Target managed table 701 and port watch is buried. It turns out that the target of a line with which **** of the Target managed table 702 and port watch were buried as a result is RCU. Although there is what of three lines has the same **** by drawing 7, this shows that a remote copy of the port of three different numbers of the same RCU (that is, it is also the same storage) is possible.

[0028] In an actual system, the path information (here, a logical path (logical page number) is called) for performing data communication is clearly set up and managed on application level between RCU(s) which are MCU and the target which are an initiator, and the logical-page-number (logical path) managed table 801 of drawing 8 is used as a table of the information which carries out setup reference by the user as an example here.

[0029] By drawing 8, the flow of the information table of the logical path set as MCU with a user for convenience is made into an application layer, and, on the other hand, the RCU information and logical-path information flow which are decided on a fiber channel are explained as a fiber channel layer. The information on the target of remote ** (****, port watch) is set to the logical-page-number managed table 801 in advance. Avail of effective logical page number (logical path) is set to 01, and Avail of logical page number without target information is set to 00. Moreover, the state of 80 and decision is set to 00 for an undecided state in a fiber channel layer as Status of logical page number.

[0030] On the other hand, in the fiber channel layer, the destination address on a fiber channel, **** returned to process login, and port watch are stored in the Target managed table 802. If RCU is decided, it will refer to the logical-page-number managed table of an application layer, only logical page number of **** which exists among [both] both tables, and port watch will be confirmed, and the logical page number-Target translation table 804 will be updated. Status of logical page number which only logical page number decided by the logical page number-Target translation table 804 became an actually usable logical path, was made to reflect the definite information on logical page number also in the logical-page-number managed table 803 of an application layer by the opportunity by which the logical page number-Target translation table 804 is updated, and had been decided on the logical page number-Target managed table 804 is updated from 80 to 00. The storage and the port of a remote copy place are pinpointed by the above. It means that a remote copy is carried out in three target ports from specified MCU which is this example.

[0031] participation of a node also with the new link established at once by the fiber channel or according to the event of removing a node from a link, although it is necessary to reestablish a link again, it starts with the above-mentioned login procedure also in this case. The justification of logical page number is guaranteed by table management shown by drawing 7 and drawing 8 also in that time.

[0032]

[Effect of the Invention] If this invention is used as shown in the above example, the remote copy between storage will become realizable through a fiber channel. And a merit called the transfer performance and long-distance connection which a fiber channel interface has by fiber channel connection can be harnessed in a remote copy and host I/O.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. *** shows the word which can not be translated.

3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The system configuration view of the environmental construction in this example

[Drawing 2] The frame format used for the example of this invention, and detail drawing of the data field

[Drawing 3] Drawing showing the login process used for the example of this invention

[Drawing 4] Drawing showing the login process in the example of this invention

[Drawing 5] It is the control flow chart of a login transmitting side to the example of this invention.

[Drawing 6] It is the control flow chart of a login receiving side to the example of this invention.

[Drawing 7] The table showing the flow of the target port detection in the example of this invention

[Drawing 8] The table control chart showing establishment of the logical path in the example of this invention

[Description of Notations]

101-102 Host computer

103-104 Fiber channel port (host)

105 Fiber Channel

106 Fiber Channel Port (Storage : for Initiators)

107-109 Fiber channel port (storage : for targets)

110 MCU

111 RCU

112 M-VOL

113 R-VOL

114 to 115 storage

201 Frame

202 SOF(Start Of Frame)

203 Frame Header

204 Data Field

205 CRC

206 EOF

207 Detailed Structure of PLOGI

208 N_Port_Name

209 Node_Name

210 Detailed Structure of PRLI

211 Originator process assiator

212 Responder process assiator

[Translation done.]

(51) Int.Cl.	識別記号	F I	テマコード* (参考)
G 0 6 F 13/00	3 0 1	G 0 6 F 13/00	3 0 1 R 5 B 0 1 4
3/06	3 0 4	3/06	3 0 4 F 5 B 0 6 5
13/12	3 3 0	13/12	3 3 0 Z 5 B 0 8 3
H 0 4 L 13/08		H 0 4 L 13/08	5 K 0 3 4

審査請求 未請求 請求項の数14 O L (全 13 頁)

(21) 出願番号	特願2000-128970(P2000-128970)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目6番地
(22) 出願日	平成12年4月25日 (2000.4.25)	(72) 発明者	小林 直孝 神奈川県小田原市国府津2880番地 株式会 社日立製作所ストレージシステム事業部内
		(72) 発明者	阿部井 大 神奈川県小田原市国府津2880番地 株式会 社日立製作所ストレージシステム事業部内
		(74) 代理人	100068504 弁理士 小川 勝男 (外2名)

最終頁に続く

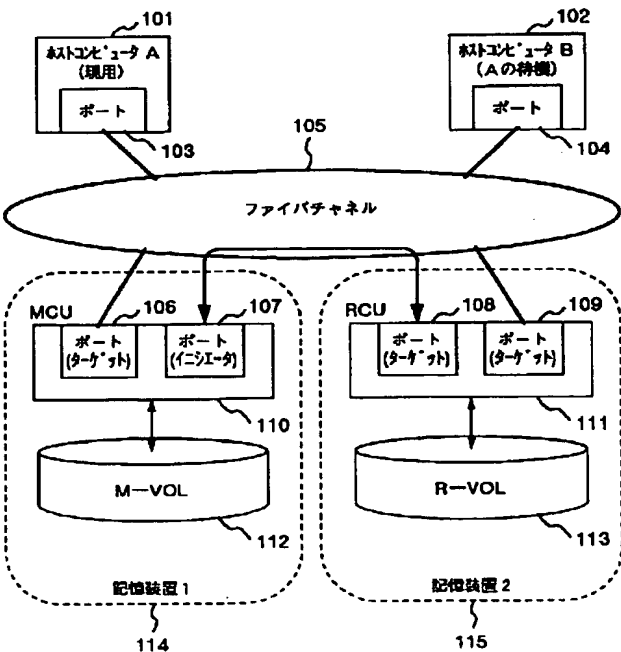
(54) 【発明の名称】 記憶装置のリモートコピーシステム

(57) 【要約】

【課題】 記憶装置間のリモートコピーを実現するコンピュータシステムにおいて、記憶装置のリモートコピー転送の速度が遅く、更に、そのためホスト I / O のスループットの妨げにもなっていた。

【解決手段】 ホストコンピュータと複数の記憶装置をファイバチャネルで接続し、ファイバチャネルを介してリモートコピーを実施する。そのために、コピー元となる記憶装置は記憶装置からのログインであることを識別できる情報を付与してログインすると共に、ログインの受領先の記憶装置は記憶装置からのログインのときのみリモートコピーできるポートを特定する情報を返すことによりリモートサイトの記憶装置及びポートの候補を決める。

図 1



【特許請求の範囲】

【請求項 1】 ホストコンピュータ及びリモートコピーのイニシエータとなるポートを持つ記憶装置とターゲットポートを持つ記憶装置がファイバチャネルを介して接続され、前記イニシエータとなるポートを持つ記憶装置は自身が記憶装置であることを識別できる第 1 の情報を付加したログインを前記ターゲットポートを持つ記憶装置のターゲットポートに行なう制御装置を有し、前記ターゲットポートを持つ記憶装置は前記ログインの前記第 1 の情報によりログイン元が記憶装置であることを識別したとき、前記ターゲットポートを特定するハードウェアに固有で且つ不変な第 2 の情報を付加した応答を返す制御装置を有することを特徴とする記憶装置のリモートコピーシステム。

【請求項 2】 前記第 2 の情報はターゲットポートを持つ記憶装置のハードウェアに固有な識別子とポート番であることを特徴とする請求項 1 記載の記憶装置のリモートコピーシステム。

【請求項 3】 前記第 2 の情報はターゲットポートを持つ記憶装置の製番とポート番であることを特徴とする請求項 1 記載の記憶装置のリモートコピーシステム。

【請求項 4】 前記第 1 の情報はイニシエータとなるポートを持つ記憶装置のハードウェアに固有な識別子とポート番であることを特徴とする請求項 1 記載の記憶装置のリモートコピーシステム。

【請求項 5】 前記第 1 の情報はイニシエータとなるポートを持つ記憶装置の製番とポート番であることを特徴とする請求項 1 記載の記憶装置のリモートコピーシステム。

【請求項 6】 前記ポート番はイニシエータとなる記憶装置で保持するポートの装着位置を示すロケーションの識別番号であることを特徴とする請求項 3, 4, 5 の何れか記載の記憶装置のリモートコピーシステム。

【請求項 7】 ホストコンピュータとリモートコピーのコピー元であってマスタコントロールユニットを有する記憶装置とターゲットポートを持つ記憶装置とがファイバチャネルを介して接続され、前記マスタコントロールユニットはイニシエータポートからターゲットポートに対しマスタコントロールユニットからのログインであることを示す情報を付加してログインし、前記ターゲットポートを持つ記憶装置はマスタコントロールユニットからのログインであることを認識すると自身がリモートコピーの対象となる記憶装置として前記ターゲットポートを特定するハードウェアに固有で且つ不変な情報を付加した応答を前記イニシエータに返すことを特徴とする記憶装置のリモートコピーシステム。

【請求項 8】 前記マスタコントロールユニットからのログインであることを示す情報は前記マスタコントロールユニットに固有な識別子とポート番であることを特徴とする請求項 7 記載の記憶装置のリモートコピーシステム。

ム。

【請求項 9】 前記ポート番は前記マスタコントロールユニットで保持するポートの装着位置を示すロケーションの識別情報であることを特徴とする請求項 8 記載の記憶装置のリモートコピーシステム。

【請求項 10】 前記マスタコントロールユニットからのログインであることを示す情報はファイバチャネル通信プロトコルで使用される識別子以外の識別子であることを特徴とする請求項 7 記載の記憶装置のリモートコピーシステム。

【請求項 11】 ホストコンピュータとそれぞれコントロールユニットを持つ複数の記憶装置がファイバチャネルを介して接続され、リモートコピーのコピー元となる記憶装置のコントロールユニットは当該コントロールユニットに固有な識別子とポート番を付して他の前記記憶装置のポートにログインし、前記他の記憶装置のコントロールユニットは前記ログインのデータにリモートコピーのコピー元となる記憶装置のコントロールユニットに固有な識別番号とポート番があることに応答して当該コントロールユニットに固有な識別子とポート番を付加した応答を返すことを特徴とする記憶装置のリモートコピーシステム。

【請求項 12】 ホストコンピュータとそれぞれコントロールユニットを持つ複数の記憶装置がファイバチャネルを介して接続され、記憶装置からのログインであることを判別できる情報を付して他の前記記憶装置のポートにログインするリモートコピーのコピー元となる記憶装置のコントロールユニットと、前記ログインのデータに前記記憶装置からのログインであることを判別できる情報があることに応答して当該コントロールユニットに固有な識別子とポート番を付加した応答を返す前記他の記憶装置のコントロールユニットと、応答に含まれる前記コントロールユニットに固有な識別子とポート番とそのポートのあて先アドレスとを格納するターゲット管理テーブルと、アプリケーション層で指定されるコピー先のコントロールユニットに固有な識別子とポート番を格納する論理バス管理テーブルと、前記ターゲット管理テーブルと前記論理バス管理テーブルに共通に前記コントロールユニットに固有な識別子とポート番を有するターゲットポートをコピー先のポートとして格納する論理バスターゲット変換テーブルとを備えたことを特徴とする記憶装置のリモートコピーシステム。

【請求項 13】 前記コピー元となる記憶装置のコントロールユニットに固有なポート番とは当該記憶装置のコントロールユニットで保持するポートの装着位置を示すロケーションの識別情報であることを特徴とする請求項 11, 12 の何れか記載の記憶装置のリモートコピーシステム。

【請求項 14】 前記コピー元とは別の他の記憶装置のコントロールユニットに固有なポート番号とは、当該記憶

装置のコントロールユニットで保持するポートの装着位置を示すロケーションの識別情報であることを特徴とする請求項11、12の何れか記載の記憶装置のリモートコピーシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、コンピュータシステムの記憶装置間でのリモートコピーに関する。

【0002】

【従来の技術】 コンピュータシステムのディスク記憶装置などの記憶装置のボリューム（記憶データ）のコピーを遠隔地にある別の記憶装置に作成することが行なわれている。これは、地震などの災害あるいは何らかの重大な障害の発生によりコピー元の記憶装置のデータが読み出せない事態になったときに対処するものであり、このときはコピー先のデータがコピー元のデータに代わって使用される。このようなコピーを作成すること、または作成されたコピーをリモートコピーと呼ぶ。リモートコピーは特に高い信頼性が要求されるシステム、例えば銀行システムなどに採用されている。

【0003】 従来の記憶装置間でのリモートコピーにはESCON技術を駆使したものが存在する。特開平6-236340号公報がその公知例として挙げられる。

【0004】 ところで、ホストコンピュータと記憶装置との間の情報の転送には近年高速転送用チャネルインタフェースとしてファイバチャネルが使われている。このようなファイバチャネルのプロトコルではファイバチャネルを介して点在する個々の装置（ノード）または、そのポートを他と識別する為には、そのプロトコル規格に基づいたユニークな情報が使われる。WWN(World Wide Name)、AL_PA(Arbitrated Loop Physical Address)等がそうである。

【0005】 データ通信手段にファイバチャネルを使うと、伝送速度が速い(最大100MB/S)、ケーブルを長く延ばせる(単体ケーブルで最大10km、ファブリック接続では無限)といった性能と距離の2つの大きなメリットがある。前記記憶装置においてもファイバチャネルを使ったホストコンピュータとのシステム構成ではこのメリットが得られる。

【0006】

【発明が解決しようとする課題】 ところで、リモートコピーを実施するシステムにおいてはローカルの記憶装置からリモートの記憶装置へのデータ転送を要する為、リモートコピーを実施しない場合と比較すると、ホストI/O(ホストと記憶装置との間のデータの入出力)のスループットは確実に落ちる。それは、リモートコピーの場合、記憶装置が装置内部に持つデータキャッシュ上に保持されたライトデータが通常より長くデータキャッシュ上に残される為であり、データキャッシュの使用効率が悪いからである。ライトデータが長くデータキャッシ

ュに残される原因はリモートコピーの転送性能に起因している。前述の通り、従来までのリモートコピーの遠隔接続にはESCONが用いられており、ホストI/Oの転送速度に比べ、リモートコピーの転送速度は非常に遅く、結果的にホストコンピュータと記憶装置間の転送手段にファイバチャネルを用いても、リモートコピーを行うが為にホストI/Oの転送性能を減少させてしまう。

【0007】 もう一点、ESCONでは遠距離接続(10km超)において、エクステンダーを経由することによりさらにケーブルを延ばすことが可能であるが、ESCON系の機器は非常に高価である。その点ファイバチャネルでファブリック・スイッチ、或いはハブを経由した場合でも、ESCON系の機器に比べれば非常に安価である為、システム構築上のコスト面においてもその高価は格段に大きい。

【0008】

【課題を解決するための手段】 以上のことから、本発明はリモートコピーのデータ転送にファイバチャネルを用いることで、ホストI/Oのスループットの妨げを最小限にしようというものである。上記を実現する為に、本発明では転送元となる記憶装置上にファイバチャネルを介して通信可能なイニシエータポートを備え、リモートサイトの記憶装置にはファイバチャネルに接続されるターゲットポートを備えており、ファイバチャネルでのリンク確立のためのファイバチャネルプロトコルを介してデータ転送が行なわれる。更に、イニシエータポートからは自装置が記憶装置であることを認識できる情報を付してターゲットポートにログインし、ターゲットポートはログイン発行元が記憶装置であることを認識したとき自ポートを特定する不変のハードウェア情報を付してイニシエータポートに返す。

【0009】

【発明の実施の形態】 以下、本発明の実施例を図面を使って説明する。図1は、本発明の実施例となるリモートコピーシステムの構成を示したものである。ホストコンピュータ101~102はいずれもファイバチャネル・インタフェース用のポート(それぞれ103~104)を有し、ファイバチャネル・インタフェースを介して記憶装置114~115と物理的に接続されている。記憶装置114~115もホストコンピュータ101~102同様にファイバチャネル・インタフェース用のポート(それぞれ106~107、108~109)を有しており、ファイバチャネル・プロトコルによる通信が可能である。ホストコンピュータ101~102と記憶装置114~115間のファイバチャネルインタフェースの接続形態(トポロジ)には、Point-to-Pointや、アービトレーション・ループ接続、ファブリック接続等、いくつかの種類が存在するが、本発明はその接続形態には依存しないため、単にファイバチャネル105と記述する。

【0010】 本リモートコピーシステムでは記憶装置1

14が有するデータボリューム(M-VOL112)をファイバチャネル105を介し、リモートの記憶装置115が有するデータボリューム(R-VOL113)にコピーすることを目的としている。その際、マスターとなる記憶装置114のホストインタフェース制御装置はイニシエータとなるポート107を有し、本制御装置をマスター・コントロール・ユニット(MCU)110と呼び、対するリモートコピー先のターゲットとなるポート108を有する制御装置をリモート・コントロール・ユニット(RCU)111と呼ぶ。本発明では、MCU-RCU間で使用するインタフェース・プロトコルにファイバチャネルを使用することを第一の特徴に挙げている。そこで、そのプロトコルの概要を説明することにする。

【0011】ファイバチャネルは独自のコマンドセットをもたない、シリアル転送方式をもつプロトコルであり、情報を非同期に送るために伝送媒体の帯域幅を有効に利用できる特色を持っている。独自のコマンドセットを持たないかわりに、ファイバチャネルを、従来のSCSI、ESCON、HIPPI、IP-3、IP等といったコマンドセットのインフラとして使用することにより、従来のプロトコル資産を継承しながら、より高速で信頼性の高い多彩なデータ転送が可能である。

【0012】ファイバチャネルはチャネルとネットワークの特徴を併せ持つインタフェースである。ファイバチャネルでは一度、転送元と転送先が確定すれば、遅延が少ない高速な転送を実現できるが、これはチャネルの最大の特徴の一つである。また、通信を希望する機器は、任意の契機でファイバチャネルの通信系に参加し、通信の目的となる相手の機器と相互に通信に関する取り決め情報を交換し、通信を開始することができるが、これはネットワークの特徴である。ここで述べた相手機器との通信に関する取り決め情報交換の手続きをログインと呼ぶ。

【0013】ファイバチャネルのインタフェースを持つ機器(例えばホストコンピュータや記憶装置)をノードと呼び、実際のインタフェースにあたる物理的な口をポートと呼ぶ。ノードは一つ以上のポートを持つことが可能である。ファイバチャネルの系全体に同時に参加できるノード数は、例えば、最大で24ビットのアドレス数、すなわち約1677万個である。一般にファイバチャネル系はいくつものループ状のファイバチャネルからなっており、ループ間を接続し、情報を中継するハードウェアをファブリックと呼ぶ。実際には、送信元および送信先のポートは、ファブリックを意識せずに互いのポートに関する情報のみを考慮して動作すればよい。図1でも簡単化して表している。各ノードおよびポートには、標準化団体(IEEE)から一定のルールによって割り当てられる、世界中でユニークな識別子が記憶されている。これは従来からTCP/IPなどで馴染みのMACアドレスに相当

するものであり、ハードウェア的に固定なアドレスである。このアドレスにはN_Port_Name、Node_Nameの2種類があり、N_Port_Nameはポート固有の値(ハードウェア・アドレス)であり、Node_Nameはノードに固有の値(ハードウェア・アドレス)である。これらは、いずれも世界中でユニークな値であることから、ノードまたは、ポートを一意に識別できるアドレスとして、WWN(World Wide Name)と呼ばれる。

【0014】ファイバチャネルでは、通信はOrdered Setと呼ばれる信号レベルの情報と、フレームと呼ばれる固定のフォーマットを持った論理的な情報とで行われる。図2はフレームの構造を示している。フレーム201は、フレームの始まりを示すSOF(Start of Frame)202と呼ばれる4バイトの識別子、リンク動作の制御やフレームの特徴づけを行う24バイトのフレームヘッダ203、実際に転送される目的となるデータ部分であるデータフィールド204、4バイトの巡回冗長コード(CRC)205、フレームの終わりを示すEOF(End of Frame)206と呼ばれる4バイトの識別子からなる。データフィールド204は、0~2112バイトの間で可変である。

【0015】次にファイバチャネル・プロトコルに基づく、送信元の機器と送信先の機器に関して互いに情報を交換するログイン手続きについて述べる。ログイン手続きの際、必須となるPLOGI(ポートログイン)フレームおよびPRLI(プロセスログイン)フレームの構造について説明する。データフィールド204のPLOGIの詳細構造207において、先頭から21バイト目~29バイト目までの8バイトの領域がN_Port_Name208を格納する領域であり、先頭から30バイト目~38バイト目までの8バイトの領域がNode_Name209を格納するものである。また、PRLIの詳細構造210において、先頭から8バイト目~11バイト目までの4バイト領域がオリジネータープロセスアソシエーター(Originator process associator)211のパラメータを格納する領域である。これは、ノードをイニシエーターとして使うとき有効となるパラメータエリアである。また、先頭から12バイト目~15バイト目までの4バイト領域がレスポンスアソシエーター(Responder process associator)212のパラメータを格納する領域である。これはノードをターゲットとして使う場合有効となるパラメータエリアである。

【0016】図3は、送信元(ログイン要求元)301と送信先(ログイン受信先)302との間に取り交わされる情報のやりとりを示したものである。ファイバチャネルのログイン手続きには数種類存在するが、ここではクラス3のログインに関して述べる。クラス3とはファイバチャネルの通信手順のタイプの一つを示すもので本発明の適用が特にこれに限られるものではない。説明の都合上例示したものである。

【0017】ログイン要求元はPLOGIフレーム303をログイン受信先へ送信する。このフレームには、ログイン要求元のN_Port_Name、Node_Name、およびその他の情報が含まれている。受信先の装置では、このフレームに含まれている情報を取り出し、ログインを承認する場合はACC304と呼ばれるフレームをログイン要求元に対して送信する。一方、ログインを拒絶する場合は、LS_RJT305と呼ばれるフレームをログイン要求元に送信する。ログインを承認するか拒絶するかはターゲット側が持つファイバチャネルのルールに照らして決められるものであり、本発明には直接関係しない。ログイン要求元は、自らが送信したPLOGIフレームに対してACCフレームの応答を検出すると、ログインが成功したことを知る。ポート間のログインが成功すると、続いてプロセスレベルでのやりとりが行われる。ログイン要求元はPRLIフレーム306をログイン受信先へ送信する。このフレームには2ポートに関連するプロセスグループ間のFCP（ファイバチャネルプロトコル）レベルの動作環境を取り交わす情報が含まれている。受信先の装置ではPLOGI受領時と同様に、ログインを承認する場合はACC307、或いは拒絶する場合はLS_RJTフレーム308をログイン要求元に対して送信する。ここまでのログイン手続きが成功するとデータ転送などのI/Oプロセスを開始できる状態となる。以上がファイバチャネルを使ったときのポート間の一般の手順である。

【0018】以下にファイバチャネルを利用したリモートコピーについて説明する。図4において、MCU401～RCU402間でのファイバチャネルを介したやり取りを示す。前述の通り、MCU401～RCU402ではファイバチャネル上の経路を確立するためのログイン手続きを行う。まず、MCU401からPLOGIが送信され、RCU402ではPLOGIの承認を通知するACCを送信する。これはホストからPLOGIが送られ、それを受けたターゲットデバイスがACC応答する場合と何ら変わりはない。続いて、MCU401はPRLIを送信するが、本発明では当該記憶装置がMCU401であることをRCUへ知らせる為の手段としてPRLIフレームのペイロード（データフィールドの実データ）403の2W目（Originator Process associatorのパラメータ領域）にMCUの製番（MCUを備えた記憶装置の製番ともいえる）とポート番を載せて送信する。PRLIフレームを受け取ったRCU402は、フレームのペイロードの2W目を切り出し、MCUの製番及びポート番を取り出す。これによりRCUは当該イニシエータがMCUであることを識別できる。製番とはハードウェアにユニークな（ハードウェアに固有な）識別番号でありかつファイバチャネル系の変更やパッケージの入れ替えなどに係わらず不変なものである。製番と言う通常ソフトでは認識しないものを送信データ中に埋め込んで認識可能にしてある。ここでは、記憶装置であることが判る固有の識別情報であるという意味合いを持つ。本実施例

では製番を利用したがほかに記憶装置であることを認識できる情報であれば他の情報、例えばファイバチャネルで使用されている規格上の識別子（ファイバチャネルの通信プロトコルで使用される識別子）以外の識別子であっても良い。このように、RCUは製番とポート番の有無でMCUからのログインか、ホストコンピュータからのログインかを判別する。MCUポート番は1つの記憶装置の中で固有のもので且つ不変なものである。

【0019】ポート番とは、記憶装置に具備されるポートの記憶装置内の装着位置を示す識別情報であり、1つの記憶装置内で固有なロケーション番号として定義される。例えば、最大32ポート具備可能な記憶装置であれば、1から32までの番号を付与しても良い。このポート番はロケーションを示す情報であるので例えば、別のN_Port_Nameを持つポート基板を装着しても不変である。一方、N_Port_Nameで示される識別情報は各ポートに付与されたアドレスであり、ポートを構成するハードウェア（例えばポート基板）に依存する。したがって、ポート基板が故障し、別のポート基板を装着すると当該ポートのアドレスは変更される。

【0020】PRLIを送信してきたイニシエータがMCUであると判明したら、RCUはACCフレームのペイロード404の3W目（Responder Process associatorのパラメータ領域）にRCU製番とRCUポート番を載せて応答する。ACCフレームを受け取ったMCUは、フレームのペイロードの3W目を切り出し、RCUの製番及びポート番を取り出す。これによりMCUはログインした相手がリモートコピーをすることのできる記憶装置の一つであることが認識できると共にRCUのポートを特定できる。このように通常のログイン手続きにターゲットポートを識別する方式を盛込んだ処理の流れを、図5、図6では要求元と受領先とに分けたフローチャートで詳細に解説してある。

【0021】図5の要求元の処理フローに示すようにMCUが行うログイン手続きでは最終的にファイバチャネル上に点在するRCUだけを抽出することを目的としている。PLOGIが成功後のPRLIに対するACC応答でACCフレームから製番、ポート番を取出すことが出来たターゲットポートのみリモートコピーの対象として意味を持ち、それ以外のターゲットポートに対してはLOGO（ログアウト）を発行してログアウト処理を行う。このログイン手続きをファイバチャネル上のすべてのターゲットポートに対して行い、製番、ポート番を取出せたターゲットポートが実際にリモートコピーを行うペアを形成するRCUの候補となる（実際にどの記憶装置からどの記憶装置へリモートコピーを行なうかは上位のアプリケーションプログラムの指定による）。

【0022】図5を順を追って説明する。図3に示したようにあるターゲットポートに対しPLOGIを発行する（501）。PLOGIの発行が成功かどうかを判定

する(502)。即ち、LS RJTを受領すればログイン失敗処理に移る。ACCを受領すれば、それに含まれるログイン先のAL_PA、WWNを取得し相手先を特定する(503)。次に、PRLIを発行する(504)。そして、PRLI発行成功かを判断する(505)。LS RJTを受領すればログイン失敗処理に移る。ACCを受領するとそれに含まれるペイロードから製番とポート番を取り出す処理を行なう(506)。次に、取り出しが成功かを判断する(507)。取り出しが成功したと言うことは相手がリモートコピーできる記憶装置であると言うことであるのでリモートコピー先の候補として挙げる。取り出しができないと言うことは相手が記憶装置以外であることを示すからログアウト処理をする(508)。これを、前述のように総てのターゲットポートに対して連続して繰り返す。

【0023】これに対して図6はログインの受領先から見た処理フローである。受領先では要求元が送信してくるログインフレームを受領した際、そのフレーム内容をチェックしてはログインを承認するか拒絶するかを判定する。本RCUでは要求元のイニシエータがMCUである場合のみ、自身がRCUであることをMCUに通知し、そうでない(通常はホストからのログイン要求の場合)場合は単なるターゲットポートとして応答する。その為の処理がPRLIフレームからMCU製番、ポート番を取出す処理であり、取り出した場合のみ自身(RCU)の製番、ポート番をACCフレームに載せて応答する。このログイン応答をファイバチャネル上のすべてのイニシエータポートに対して行い、製番、ポート番を送り返したイニシエータが実際にリモートコピーを行うペアを形成することが可能なMCUとなる。

【0024】図6を順を追って説明する。受領先ではPLOGIを受領する(601)。そして、そのフレームをチェックして(602)、ポートログインを承認するかどうかを判定する(603)。承認しない場合は、LS RJTを送信して(608)ログイン失敗処理に移る。承認する場合はACCを作成し送信する(604)。PRLIを受領すると(605)、フレームチェックをし(606)、プロセスログインを承認するかどうかを判断する(607)。承認しない場合は同様にLS RJTを送信し(608)、ログイン失敗処理に移る。承認する場合はペイロードから送信元の製番、ポート番の取りだし処理を行なう(609)。取り出しに成功すれば(610)これはイニシエータがMCUであることを示しているの、ACCフレームに自製番、ポート番をセットし(611)、これを送信する(612)。取り出しに成功しなければこれはMCU以外のノードからのログインであるので製番、ポート番をセットすることなくACCを送信する(612)。

【0025】このときは通常の通信はできてもリモートコピーの対象とはならない。

【0026】以上のようにMCUが、RCUを識別する為の手段にRCU側の製番とポート番を用いたことには理由がある。ファイバチャネルでは、一意のポートを識別する為のアドレスとして先に述べたWWNがあり、すべてのフレームの識別IDとしてフレームに載るAL_PAとがANSI規格で存在する。通常、ファイバチャネル上のポート識別には、これらのアドレスを用いることは一般的である。WWNは通常のログイン(PLOGI)処理時に入手可能な情報であり、AL_PAは通常のポート初期化時に入手可能な情報である。この2種類の識別情報もファイバチャネル上ユニークではあるが、それだけでは当該ターゲットがRCUを示す情報にはなり得ないし、その値が不変である保証もない。WWNはハードに依存してしまう為、ハード交換時には変わる恐れがあり、一方のAL_PAはループ上に同じ値を持つポートが存在した場合、ループ初期化時に変わる可能性がある。その点、装置製番及びポート番はファイバチャネルには依存しない不変の情報であり、これをPRLIフレームに載せて返す論理は、RCU識別には有効かつ最適な方法と言える。本実施例では、PRLIフレームペイロード403の2ワード目、ACCフレームペイロード404の3ワード目といったように識別情報の格納位置まで特定して説明しているが、ペイロード位置を限定する理由はなく、あくまでもWWN、AL_PA等のファイバチャネル固有の情報以外でターゲットポートを特定できる方式であることを明示しているだけである。

【0027】このようにしてMCUはファイバチャネル上に点在するターゲットポートの中からRCUとなるポートを確定することが出来る。MCUではファイバチャネル上で検出したRCUを管理する為に、図7に示すテーブルを使用する。Target管理テーブル701はPLOGIにACC応答してきたターゲットのACCフレームから取り出したAL_PA、WWNを順番に登録する。続いて、PRLIにもACC応答してきたターゲットからは、ACCフレームのResponder Process associatorが有効なターゲットのみ製番及びポート番を取出し、Target管理テーブル701の製番、及びポート番の項を埋める。結果としてTarget管理テーブル702の製番、及びポート番の埋まったラインのターゲットがRCUであることが分かる。図7で製番が同じものが3行あるがこれは同一のRCU(即ち、同一の記憶装置でもある)の3つの異なる番号のポートがリモートコピー可能であることを示している。

【0028】実際のシステムでは、イニシエータであるMCUとターゲットであるRCU間で、データ通信を行う為の経路情報(ここでは論理パス(LP_N)と称す)は、アプリケーションレベルで明示的に設定、管理されるものであり、ここでは実施例としてユーザにより設定参照する情報のテーブルとして図8のLP_N(論理パス)管理テーブル801を用いる。

【0029】図8では便宜上、ユーザがあるMCUに設定した論理パスの情報テーブルの流れをアプリケーション

層とし、一方、ファイバチャネル上で確定されるRCU情報、論理バス情報の流れをファイバチャネル層として説明する。LPN管理テーブル801には事前にリモート先のターゲットの情報(製番、ポート番)を設定してある。有効なLPN(論理バス)のAvailを01とし、ターゲット情報のないLPNのAvailを00とする。また、LPNのStatusとしてファイバチャネル層で未確定の状態を80、確定の状態を00とする。

【0030】一方、ファイバチャネル層ではTarget管理テーブル802にはファイバチャネル上の宛先アドレスとプロセスログインに対して返された製番とポート番が格納されている。RCUが確定すると、アプリケーション層のLPN管理テーブルと照らし合わせ、両テーブル間で共に存在する製番、ポート番のLPNのみを有効とし、LPN-Target変換テーブル804を更新する。LPN-Target変換テーブル804で確定しているLPNだけが、実際に使用可能な論理バスとなり、LPN-Target変換テーブル804が更新される契機でLPNの確定情報をアプリケーション層のLPN管理テーブル803にも反映させ、LPN-Target管理テーブル804で確定していたLPNのStatusを80から00に更新する。以上により、リモートコピー先の記憶装置とポートとが特定される。この例ではある指定されたMCUから3箇所のターゲットポートにリモートコピーが実施されることを意味する。

【0031】ファイバチャネルでは、一度確立したリンクも新たなノードの参加や或いはリンクからノードを取り除くといった事象により、再びリンクを確立し直す必要があるが、この場合にも前述のログイン手続きから始める。その際でも図7、図8で示したテーブル管理によってLPNの正当性は保証される。

【0032】

【発明の効果】以上の実施例に示したように、本発明を用いればファイバチャネルを介しても記憶装置間のリモートコピーが実現可能となる。そして、ファイバチャネル接続によってファイバチャネルインタフェースが持つ転送性能、長距離接続といったメリットをリモートコピー及び、ホストI/Oに活かすことができる。

【図面の簡単な説明】

【図1】本実施例における環境構築のシステム構成図

【図2】本発明の実施例に用いるフレーム・フォーマッ

トおよびそのデータフィールドの詳細図

【図3】本発明の実施例に用いるログインプロセスを示す図

【図4】本発明の実施例におけるログインプロセスを示す図

【図5】本発明の実施例にログイン送信側の制御フローチャート

【図6】本発明の実施例にログイン受信側の制御フローチャート

【図7】本発明の実施例におけるターゲットポート検出の流れを示すテーブル

【図8】本発明の実施例における論理バスの確立を示すテーブル管理図

【符号の説明】

101～102 ホストコンピュータ

103～104 ファイバチャネル・ポート(ホスト)

105 ファイバチャネル

106 ファイバチャネル・ポート(記憶装置：イニシエータ用)

107～109 ファイバチャネル・ポート(記憶装置：ターゲット用)

110 MCU

111 RCU

112 M-VOL

113 R-VOL

114～115 記憶装置

201 フレーム

202 SOF(Start Of Frame)

203 フレームヘッダ

204 データフィールド

205 CRC

206 EOF

207 PLOGIの詳細構造

208 N_Port_Name

209 Node_Name

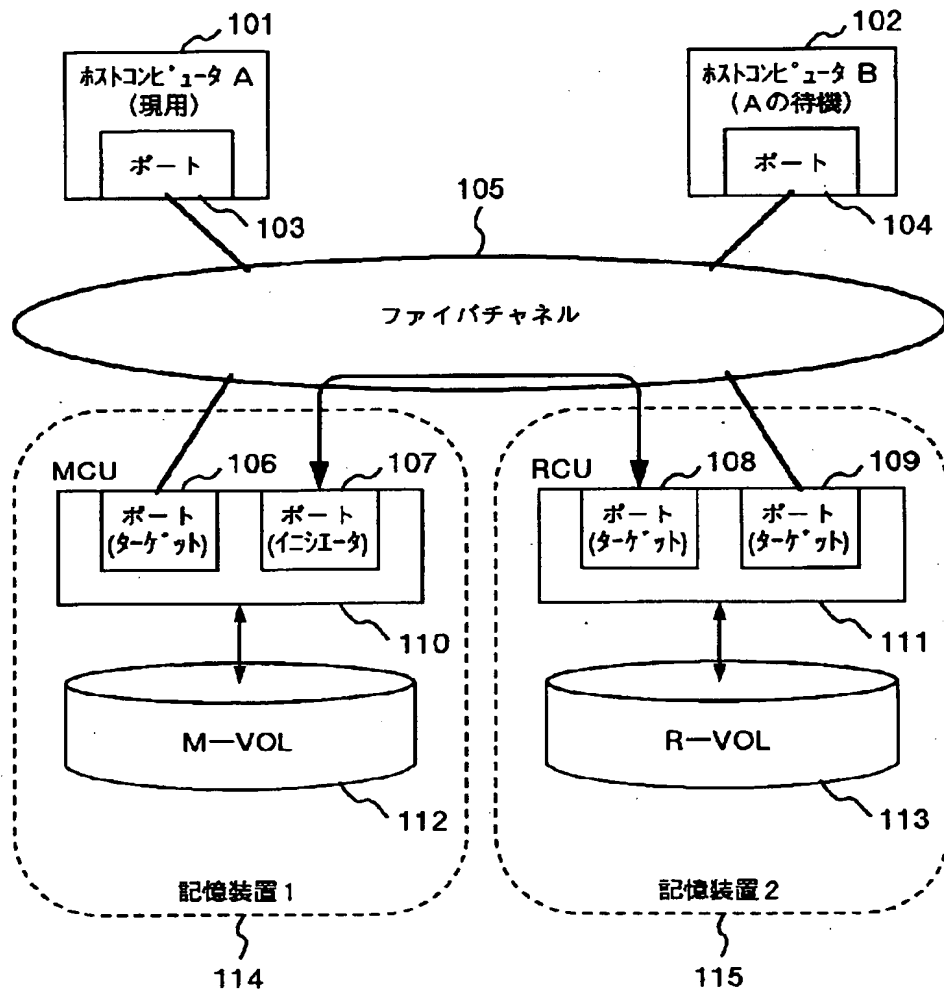
210 PRLIの詳細構造

211 Originator process assotiator

212 Responder process assotiator

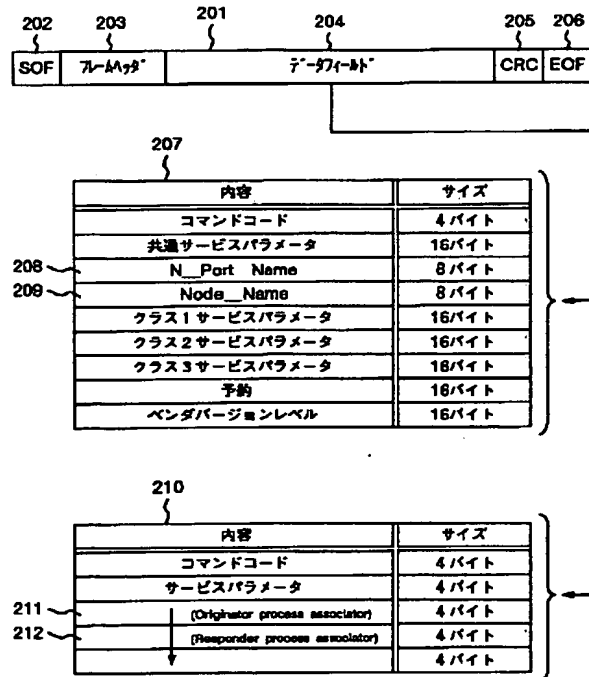
【図1】

図 1



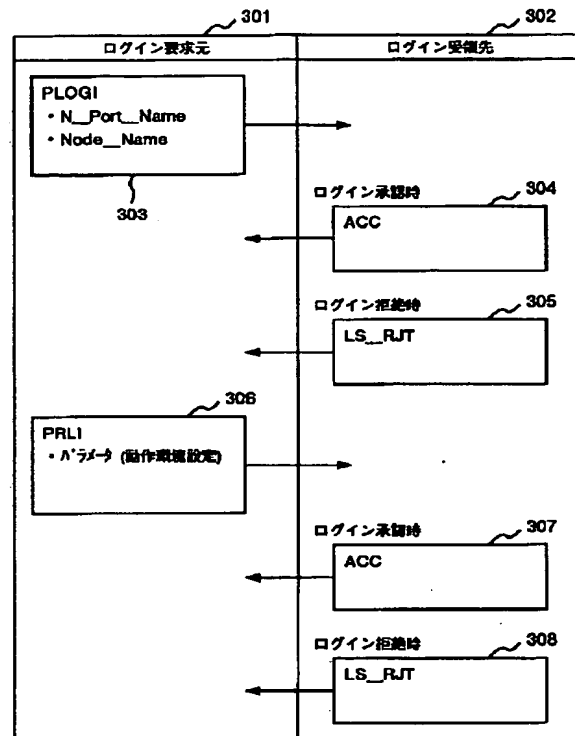
【図2】

図 2



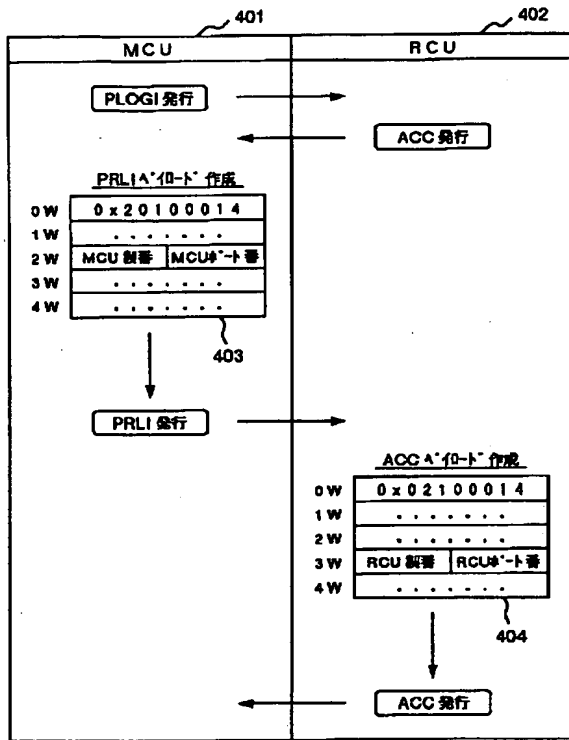
【図3】

図 3



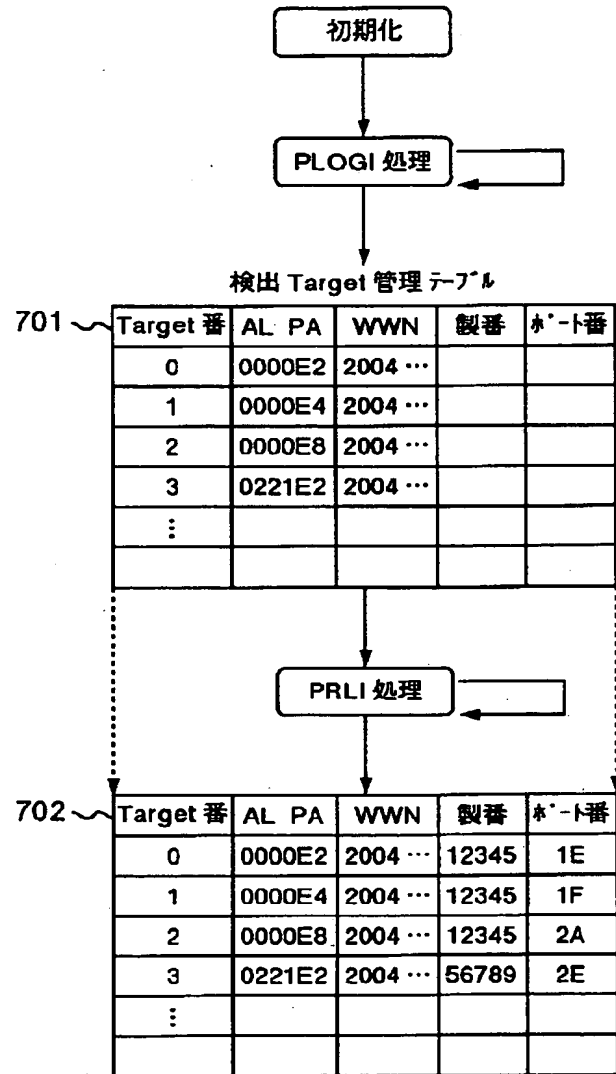
【図 4】

図 4



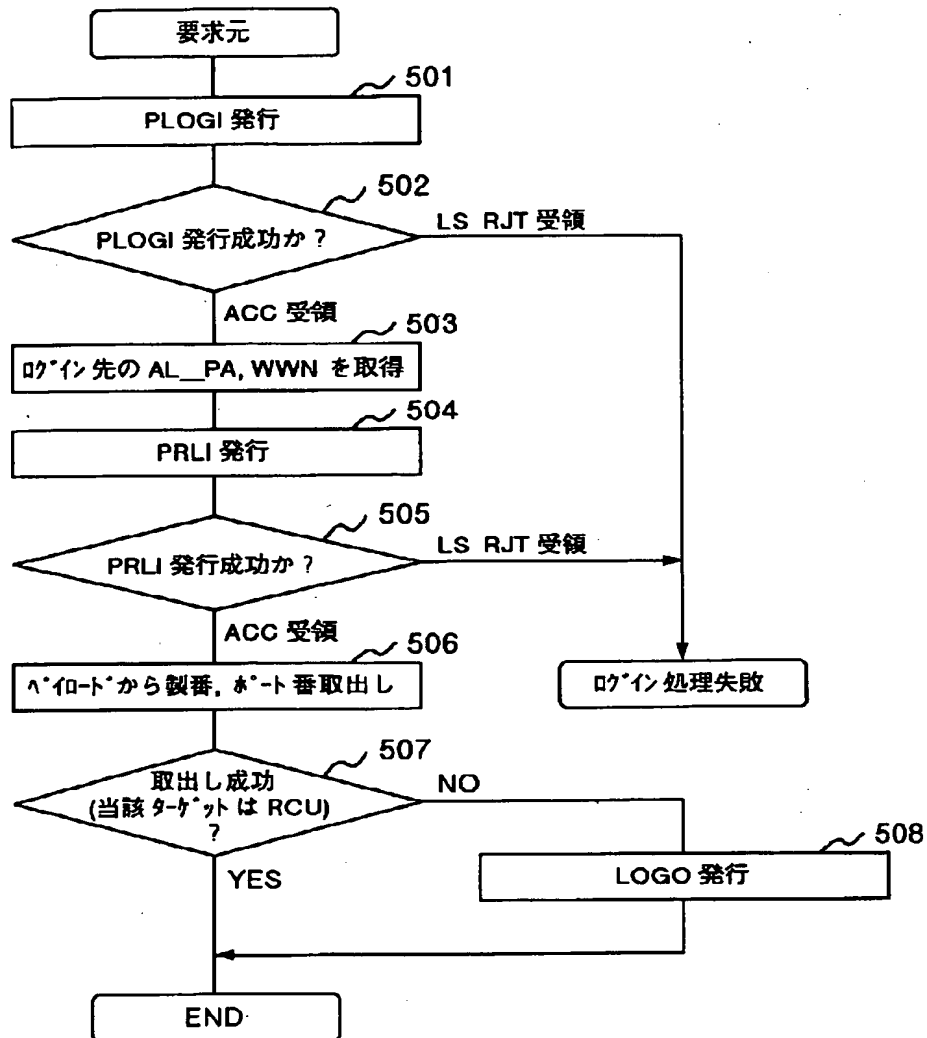
【図 7】

図 7



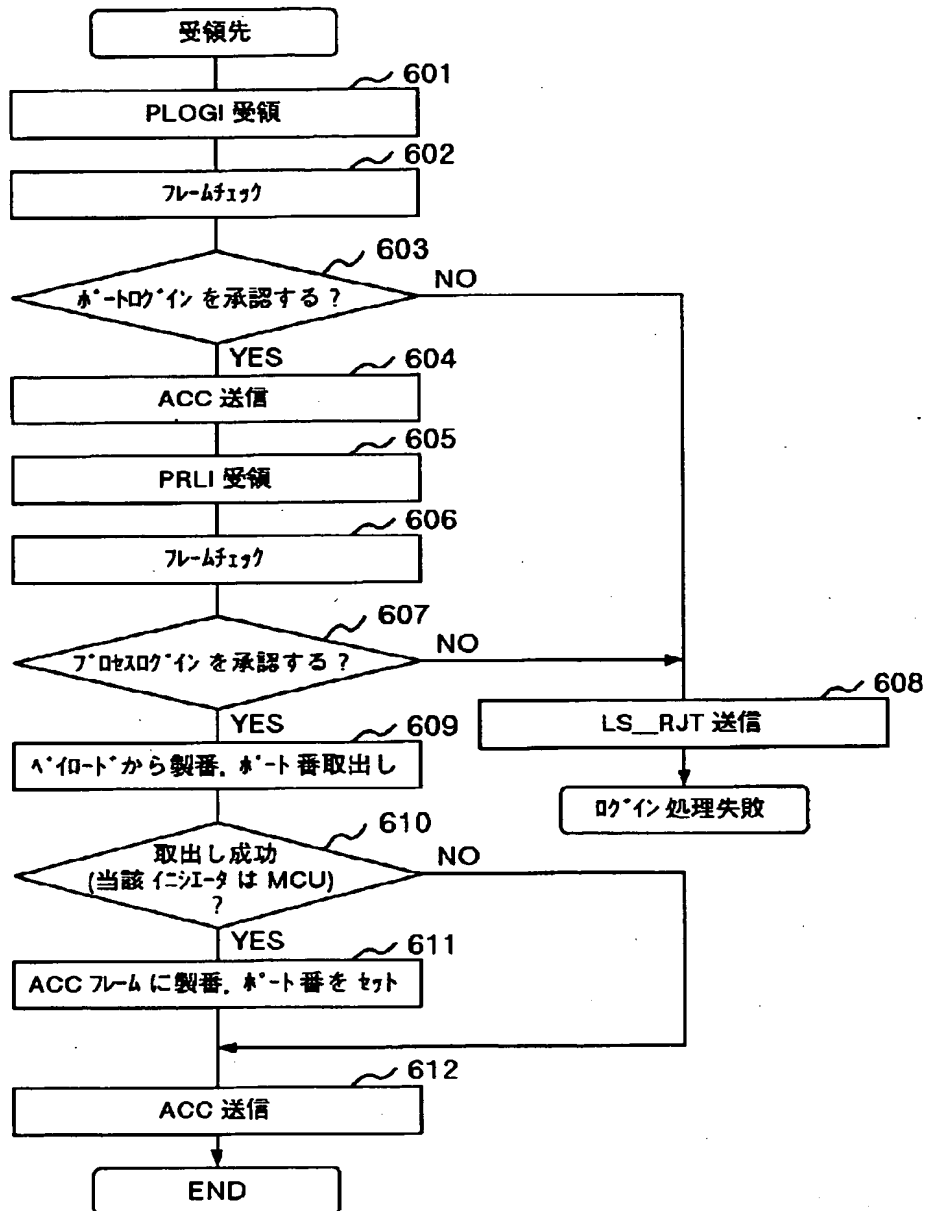
【図5】

図 5



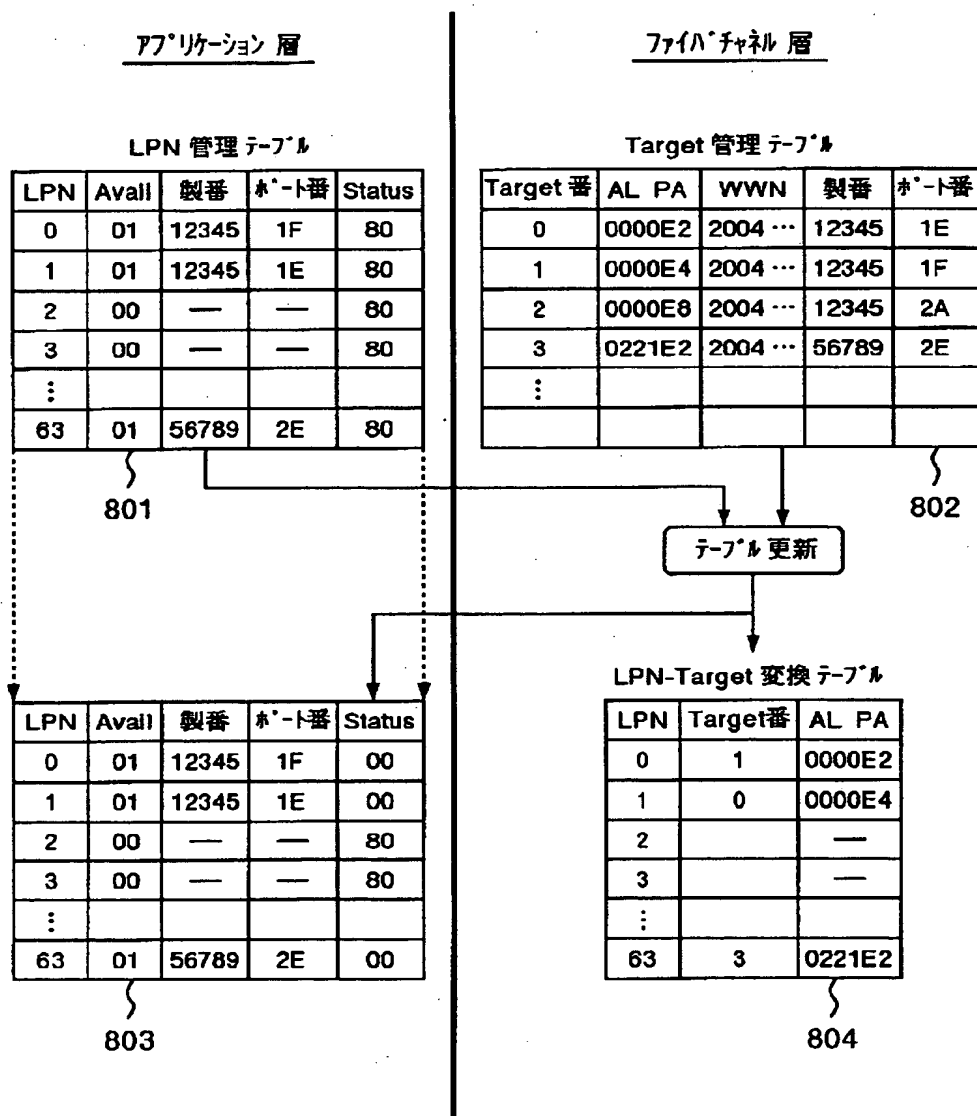
【図6】

図 6



【図8】

図 8



フロントページの続き

Fターム(参考) 5B014 EA04 EB05 GC00
 5B065 BA01 CE22 EA34
 5B083 AA01 CD13 GG08
 5K034 AA01 DD01 EE02 FF01 FF02
 HH11 HH17 HH26 NN12 NN26